DOCUMENT RESUME

ED 355 772 FL 021 025

AUTHOR Larson, Jerry W.
TITLE Computerized Adaptive Spanish Placement Test. Final Performance Report.
SPONS AGENCY Office of International Education (ED), Washington, DC.
PUB DATE [85]
CONTRACT G008440250
NOTE 16p.
PUB TYPE Reports - Evaluative/Feasibility (142)

EDRS PRICE MF01/PC01 Plus Postage.
DESCRIPTORS College Second Language Programs; *Computer Assisted Testing; Higher Education; *Language Tests; Microcomputers; Recordkeeping; *Second Languages; *Spanish; *Test Construction; Test Format; Test Reliability; Test Use; Test Validity; Undergraduate Students
IDENTIFIERS *Brigham Young University UT; *Placement Tests

ABSTRACT

A study at Brigham Young University (Utah) investigated the feasibility of computer-assisted language placement testing in higher education. Benefits and problems of this approach for test administration, individualization of item selection, and recordkeeping were examined. Four steps were followed in production of a test for Spanish placement: development of an item bank; determination of a suitable psychometric methodology; selection of a test delivery system; and validation and implementation. A 1,443-item bank was created, reduced to 1,100 items, and incorporated into 5 test forms. Tests were administered to 199 students from lower-division Spanish courses, then submitted to conventional statistical and Rasch analyses. Items were indexed to 51 difficulty levels. After final review by faculty, the 1,050 items remaining were coded, and 9 items at each difficulty level were included in the final version of the placement test. The test was then designed so the computer would present to each examinee items in the estimated range of his/her ability and inform students immediately of their probable class placement. The test is delivered on a standard IBM or compatible personal computer with two disk drives. Student records are stored and can be printed. Initial test implementation was under way at the time of the report's writing and expansion of such testing was anticipated. (MSE)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

ED355772

Final Performance Report
Jerry W Larson
Brigham Young University

Computerized Adaptive Spanish Placement Test

Grant #G008440250

Rationale for the Study

Nearly all foreign language departments in institutions of higher education are faced each semester with decisions regarding the placement of new students. Large language departments, such as the Department of Spanish and Portuguese at Brigham Young University, must place hundreds of students in lower division courses. Because of the enormity of the task--considering the time and manpower required--it seems judicious to utilize the computer to the extent possible to assist in this effort.

The purpose of this study was to investigate the possibility of using the computer as as a delivery system for Spanish placement testing, not only as an "administrator" of the test but also as the manager, or record keeper, of the test results. Additionally, in its role as test administrator, the computer was to adapt the test to the examinee: present items whose difficulty level matched the ability level of the student. This type of testing (i.e., computerized adaptive testing) has been tried in areas such as mental measurement and attitude testing but, to my knowledge, not in foreign language testing.

FL 021 025

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)
This document has been reproduced as received from the person or organization originating it.
Minor changes have been made to improve reproduction quality.
Points of view or opinions stated in this document do not necessarily represent official OERI position or policy.

BEST COPY AVAILABLE

The benefits and advantages of using a computerized adaptive language test (CALT) include the following:

1) Benefits to CALT examinees: Since CALT tests are tailored to the individual, the examinee does not have to experience the boredom or frustration commonly associated with taking lengthy tests whose items ususally range in difficulty from extremely simple to very difficult. Additionally, the examinee will spend much less time taking the test, since fewer items are required to assess the student's ability. Immediate feedback is another aspect of CALT testing that is desirable to the examinee.

2) Administrative benefits of CALT testing: Because the computer serves, to a large extent, as the test administrator, controlling item selection, sequencing, and delivery, fewer persons are needed to handle testing chores. Also, less time is needed to administer each individualized test. And, in the event that certain test items require changing or deleting, this can be done without having to reprint new test booklets.

Computerized adaptive testing eliminates the problems of cheating and test compromise, since each test is virtually unique. The headaches of test storage, handling test booklets and answer sheets (and their costs) are eliminated as well. Also alleviated

are the time and expense involved in scoring and reporting test results.

Development of a Computerized Adaptive Spanish Placement Test

Four principal steps were followed in producing the CALT test for this study: 1) development of a test item bank, 2) determination of an suitable psychometric methodology, 3) selection and development of an appropriate CALT delivery system, 4) validation and implementation of the test.

Placement test item bank. The Interim Report (submitted March, 1985) discussed issues and procedures regarding the development of the test item bank. At that time 1443 items had been written and given to members of the Spanish Department faculty for review. Eleven hundred of those items were selected and put into five separate test forms (A-E), each having two parts; these tests were administered to 199 students from lower division Spanish courses (Spanish 101, 102, 201 and 202/302). The tests were submitted to two types of statistical analysis after the students had completed them: conventional and Rasch analyses. The conventional analysis consisted of determining the Hoyt and KR-21 reliability coefficients and the point biserial coefficents. The results are presented in Table 1.

Table 1

Reliability and Point Biserial Coefficients
of Test Forms A - E

Test Form	Part	Hoyt	KR-21	PBIS
A	1	.95	.93	.79
	2	.94	.92	.82
B	1	.96	.95	.74
	2	.95	.94	.82
C	1	.97	.96	.82
	2	.97	.96	.82
D	1	.96	.95	.79
	2	.95	.94	.83
E	1	.93	.91	.77
	2	.87	.83	.85

n = 199

The Rasch analysis was run using Microscale Plus™, a statistical program written specifically for computing Rasch fit statistics. Using this program, item difficulty indices were obtained for each of the items on each of the five test forms. All of the items were "anchored" together and located along a single difficulty continuum. This enabled us to select items of specified difficulty, which is crucial to constructing an adaptive test. The range of difficulty of the items was from

-3.6 to +6.0 logits[1], which represents 96 difficulty levels. Since a limited number of items were located at the extremes of the contiuum, the 96 levels were ultimately grouped into 51 levels (0-50), covering a difficulty range of -2.0 to +5.8. This range of difficulty easily spanned the ability level of students in the lower division Spanish courses.

After the test items had been calibrated and assigned a difficulty index, they were again given to members of the Spanish Department for careful scrutiny, checking for errors in spelling, inaccuracies in grammar, misleading or tricky distractors, or any other potential problems present in the items. As a result of this review, an additional 51 items were removed from the item bank, leaving about 1050 items that could be used in the final placement test.

Each of the remaining items were coded, indicating item type (i.e., grammar, reading, or vocabulary), difficulty index, question number, and the letter of the corrected response.[2] Nine items at each of the 51 difficulty levels were included in the final version of the placement test. These test items were placed in a random access file, which was created to provide

---

[1]A logit is simply a unit of measure along the difficulty continuum.

[2]The items used in this placement test are all objectively-scored multiple-choice items focusing on either grammar analysis, reading comprehension, or vocabulary knowledge.

quick and random selection of items within individual difficulty levels.

Methodology. The test was designed so that the computer would present to each individual examinee items that are in the estimated range of his or her abililty. Placement into a specified course would be determined by the student's performance on these items. In order to avoid a student's being given the same item twice, a sequential file was created that contained a key for each question in the test item base. This file was read into an array in memory at the beginning of each individual test session. That array had a flag to indicate whether a question had been used. If the question had appeared previously during the test in progress, it would not be selected a second time.

In order for the test to "zero in" on the ability range of the student, the following procedure is followed: When a student begins the test, the computer selects an item from difficulty level 15, which is at the lower end of the ability range of most beginning students. If the examinee answers the item correctly, an item at level 25 is randomly selected and displayed on the screen; if the examinee answers the initial item incorrectly, the next item that appears will be from level 10. For the first five questions, the test will increment the difficulty level by 10 for the next question if the response is correct, or it will decrease by 5 if the response is incorrect. This allows the test to adapt quickly to the approximate ability level of the examinee. After

the first five questions have been attempted, the difficulty level adjusts by plus or minus one, depending on the response. This allows for pinpointing the examinee's actual level of performance. The test is terminated when one of the following conditions is met: 1) the examinee has five correct responses (not necessarily in succession) on any one level; 2) the examinee has four incorrect answers on any one level; or 3) the student has gone through 31 items without meeting one of the above two conditions. It was found that condition #2 was the most common reason for termination of the test. Students met condition #1 only if they were answering at the highest level of difficulty. Condition #3 was most commonly met at the lower ability levels where guessing was a greater factor.

At the conclusion of the test, the examinees are immediately informed of their probable class placement level. This is determined by the computer as it equates the student's exit level with corresponding course performance level parameters. These parameters, or cutoff points, were determined from the performance of the students from the courses in question who originally served as the norming group for the test.[3]

CALT placement test delivery system. The CALT placement test is designed for administration on a standard IBM-PC, or compatible

---

[3]These cutoff points are subject to alteration depending upon the results of subsequent validation of the placement test itself. Indeed, it appears that they might be slightly high.

microcomputer, with 64K memory. Either a monochrome or color monitor may be used. Two disk drives are required: drive A houses the program disk, and drive B records the examinee's test performance data. After the test master disk is booted, instructions are displayed, informing the student how to proceed with the test[4] and instructs him to enter his name and social security number; this information is retained for each examinee record. After the student enters this information, the computer automatically displays the first item of the test. When the student has answered a sufficient number of items at his level of ability, the test will automatically terminate, and the recommended class placement level is displayed on the screen for the student.

At the conclusion of an individual test or at the end of the day, the person in charge of the testing operation can print to the screen or on hardcopy all or selected student records. The data disk will accommodate up to 200 different student data files. (Examinees are not aware of the procedure for printing hardcopy of the test results.) The printout of each record contains the examinee's name, social security number, date the test was taken, time the test began, time the test was completed,

---

[4]The instructions to the student simply explain that he is to respond by typing the letter of the correct answer choice. All keyboard keys other than those corresponding to answer choices are disabled, thus preventing any disconcerting error messages and distractions from the task at hand.

level of performance,[5] number of right and wrong responses, percentage of correct responses, and class level placement. The printout also displays the response pattern of the examinee, indicating which items (by item identification number) were presented to the examinee, the difficulty index of each item, and whether the item was answered correctly or incorrectly. (This information may prove useful for item response analysis or for isolating particular learning deficiencies of students. Actual printouts of three of the students who took the placement test are included at the end of this report.)

Implementation of the CALT placement test. Further validation studies must be run on the test before it can be used on a full-scale basis. An initial validation of the test has been conducted using students from the courses into which the test is designed to place students. Seventy-four students ranging in ability from only a brief encounter in high school to students completing Spanish 302 have taken the test. During the final week of the semester, selected teachers were asked to send to take the placement test some of their students whom they felt were ready to move on to the next course level. It was felt that the responses and level designations of these students would provide accurate information for determining the course level

---

[5]The printout shows whether the student missed four items at the level just above his assigned level, or whether he successfully answered five items at this level.

cutoff points. As it turns out some students who were not ready to advance to the next higher course took the test as well; therefore, it will be necessary review the outcome of the testing session on a student-by-student basis before being able to accurately determine the minimum ability level required for each course. As stated earlier, it appears that the intial level designations for a couple of the courses were slightly high, (e.g., off by about two or three tenths of a logit).

Once the cutoff levels have been determined, the Department of Spanish and Portuguese will use the CALT placement test prior to the beginning of each semester to screen and place students in appropriate courses. Using 15 of the microcomputers in the Humanities Learning Resource Center at Brigham Young University, it is possible to test about 300-350 students in an eight-hour day, since the tests require on the average approximately 20 minutes to complete.[6]

Other language departments in the College of Humanities have expressed interest in developing CALT placement tests for their programs as well. The linguistics department has already completed a preliminary calibration of ESL test items and is hoping to implement fully a placement/screening process in their English-as-a-second language program during the coming semester.

---

[6]During the first validation check on the test, the time required to complete the test ranged from an average time of 18.1 minutes for beginning students to 7.5 for students in 302.

Conclusions and Recommendations

Although initial observations and evaluations of the product of this study are extremely impressive and promising, further work is needed in order to produce a test in which one would have total confidence. For example, additional research might be done to perfect the procedure for determining when an examinee's test should be terminated. A possibility might be to have a standard error calculated as the test progresses, having the test terminate when the error coefficient drops below a specified level.

Another application of the computerized adaptive testing format might well be in the areas of proficiency, achievement, and diagnostic testing. Many of the problems dealt with in developing this placement test would be similar in nature to those encountered with these other types of tests.

Given the technology of the day, it is possible to expand the scope of computerized adaptive tests by adding a listening component. After the items have been written, recorded, administered to a norming group, and calibrated, they could be accessed during the testing session using a random-access audio disk player. Audio disk technology allows for almost instantaneous retrieval of any prerecorded segments on the disk. The computer could be programmed to seek and retrieve audio segments based upon the examinee's responses.

The results of the past year's efforts to develop a Spanish computerized placement test are gratifying. We have shown that it is possible to produce a test that can substantially reduce the burden of placement testing for language departments with large enrollments in lower division courses. While further work is necessary to evaluate and improve the product of this study, most of the major obstacles have been overcome.

SPANISH COMPUTER ADAPTIVE PLACEMENT TEST REPORT

Elizabeth A. Robertson 564-65-5310 Date: 12-03-1985 Time started: 14:58:41

Current class is Spanish 201

Difficulty Level	Test Item	Student Answer	Right/ Wrong
15	129	C	r
25	214	B	w
20	173	A	r
30	260	C	r
40	354	D	r
50	439	B	w
49	432	A	w
48	427	D	r
49	435	D	r
50	446	D	w
49	429	D	w
48	428	D	w
47	416	A	r
48	423	D	r
49	430	A	w
48	426	C	r
49	434	D	w

There were four incorrect answers at level 49
LEVEL: 49 Number right: 9 Number wrong: 8
Final percentage is .5294118 .
This student placed at level 48 .
Based on this student's performance on this test, he/she should be placed in Spanish 202 or 302 if a returned missionary.

Test completed at 15:08:32

BEST COPY AVAILABLE

SPANISH COMPUTER ADAPTIVE PLACEMENT TEST REPORT

Valerie Seely 529-06-5408 Date: 12-04-1985 Time started: 10:21:03

Current class is Spanish 101

Difficulty Level	Test Item	Student Answer	Right/ Wrong
15	130	D	w
10	88	D	r
20	167	C	r
30	264	A	w
25	213	A	r
35	303	D	r
36	313	B	w
35	305	D	w
34	299	D	w
33	290	C	w
32	276	D	r
33	285	D	r
34	300	A	w
33	289	D	r
34	293	C	w
33	286	A	r
34	301	D	r
35	304	B	w
34	296	B	r
35	310	B	w
34	297	A	r
35	308	B	w

There were four incorrect answers at level 35
LEVEL: 35 . Number right: 11 Number wrong: 11
Final percentage is .5 .
This student placed at level 34 .
Based on this student's performance on this test,
he/she should be placed in Spanish 102.

Test completed at 10:43:13

BEST COPY AVAILABLE

SPANISH COMPUTER ADAPTIVE PLACEMENT TEST REPORT

John R. Doty 484-92-8181 Date: 12-05-1985 Time started: 01:48:49
Current class is Spanish 102

Difficulty Level	Test Item	Student Answer	Right/ Wrong
15	126	A	r
25	217	B	r
35	310	C	r
45	398	D	r
50	440	A	w
45	394	C	w
44	389	B	w
43	380	B	r
44	388	C	w
43	375	D	w
42	369	A	w
41	359	B	w
40	353	C	r
41	364	B	w
40	350	A	r
41	362	C	w
40	349	D	r
41	357	B	w

There were four incorrect answers at level 41
LEVEL: 41 Number right: 8 Number wrong: 10
Final percentage is .4444445 .
This student placed at level 40 .
Based on this student's performance on this test,
he/she should be placed in Spanish 201.

Test completed at 01:59:10

BEST COPY AVAILABLE